Name: Qi Li

Student ID: 1005299302

Course: INF2178-Experimental Design for Data Science

Instructor: Shion Guha

Content: Technical Assignment IV

Introduction

The dataset used in this assignment is derived from a longitudinal study focusing on MRI results of patients with and without dementia. The objective of this study is to analyze the variations in Estimated Total Intracranial Volume (eTIV) within the context of the participants' diagnostic groups and across different visits. Using Python and statistic methods, we can check how eTIV changes over time and between different diagnostic categories.

Exploratory data analysis

Figure 1 shows the kernel density estimation (KDE) plot to show the distribution. The 'eTIV' variable was represented on the x-axis. The KDE plot provided insights into the distribution of Estimated Total Intracranial Volume (eTIV) across various visits. While the overall distribution indicated no significant differences in eTIV, noticeable discrepancies were observed in the peaks of the density curves.

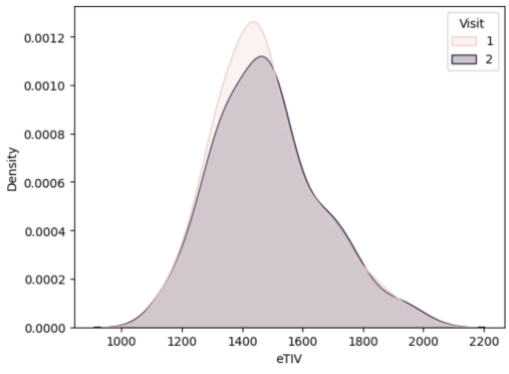

Figure 1. KDE of eTIV within Visit

In Figure 2, the boxplot visualization conducted on the dataset revealed significant differences in Estimated Total Intracranial Volume (eTIV) among different diagnostic

groups. This observation suggests potential structural differences in the brains of individuals diagnosed with dementia during the course of the study ('converted') compared to those without dementia.

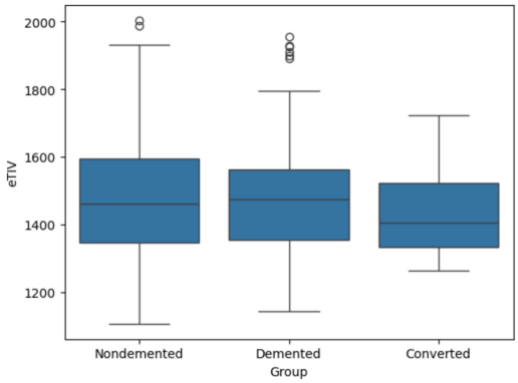

Figure 2. Boxplot of eTIV within Group

In Figure 3, the pointplot visualization conducted on the dataset yielded additional insights into the relationship between Visit, Estimated Total Intracranial Volume (eTIV), and Group. Notably, differences were observed between Visits 1 and 2, with a discernible trend towards an increase in eTIV, particularly evident in the Converted and Demented diagnostic groups. The significant differences in eTIV were primarily attributed to variations between different diagnostic groups, namely Demented, Converted, and Nondemented.

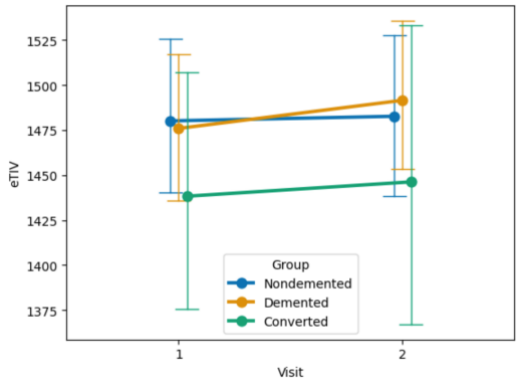

Figure 3. Pointplot of eTIV within Group and Visit

Results

The results of the repeated measures ANOVA conducted are shown in Table 1. The repeated measures ANOVA indicates a significant main effect of Visit on Estimated Total Intracranial Volume (eTIV) ($F(1, 143) = 9.246696$, $p = 0.002807$). However, there

is no further information provided regarding the effect size (ng2) or epsilon (eps) due to the absence of values for these parameters. The significant main effect of Visit suggests that there are differences in eTIV across the different visit numbers.

Source	SS	DF	MS	F	p-unc	ng2	eps
Visit	5573.92	1	5573.92	9.25	0.00	0.00	1.0
Error	86200.58	143	602.80	NaN	NaN	NaN	NaN

Table 1. The results of the repeated measures ANOVA

In Table 2, the difference in Estimated Total Intracranial Volume (eTIV) between Visit 1 and Visit 2 is statistically significant. The negative Hedges' g value of approximately -0.05 indicates a small effect size, suggesting that eTIV tends to decrease from Visit 1 to Visit 2.

Contrast	T	dof	p-unc	BF10	hedges
Visit	-3.04	143.0	0.00	7.532	-0.05

Table 2. The results of the repeated measures ANOVA

The two-way mixed-design ANOVA examines the effects of two independent variables on the dependent variable. The main effect of 'Group' is not statistically significant ($F(2, 141) = 0.297$, $p = 0.743$). This indicates that there are no significant differences in eTIV among the different diagnostic groups. The main effect of 'Visit' is statistically significant ($F(1, 141) = 9.225$, $p = 0.003$). This suggests that there are significant differences in eTIV across different visits.

Source	DF1	DF2	MS	F	p-unc	np2	eps
Group	2	141	18712.354	0.297	0.743	0.004	nan
Visit	1	141	5573.920	9.225	0.003	0.061	1.000
Interaction	2	141	502.392	0.831	0.438	0.012	nan

Table 3. The results of two-way mixed-design ANOVA

In table 4, it provides insights into the comparisons made between different diagnostic groups and visits concerning eTIV. For the 'Visit' contrast, a statistically significant difference was found between Visit 1 and Visit 2 (T = -3.041, p = 0.003), indicating a significant change in eTIV over time.

Contrast	Paired	Parametric	T	p-unc	BF10	hedges
Visit	True	True	-3.041	0.003	7.532	-0.049
Group	False	True	-0.897	0.382	0.423	-0.249
Group	False	True	-0.796	0.437	0.393	-0.210
Group	False	True	0.159	0.874	0.189	0.027
Visit * Group	False	True	-0.863	0.400	0.413	-0.235
Visit * Group	False	True	-0.845	0.409	0.406	-0.221
Visit * Group	False	True	0.034	0.973	0.187	0.006
Visit * Group	False	True	-0.927	0.367	0.432	-0.258
Visit * Group	False	True	-0.747	0.465	0.382	-0.199
Visit * Group	False	True	0.280	0.780	0.193	0.048

Table 4. The results of pairwise t-tests

Through Python, the statistical power of 0.91 while maintaining a significance level of 0.05 and detecting an effect size of 0.7, the calculated appropriate sample size for this scenario was approximately 45.45 participants.